

The new challenges of multiplex networks: measures and models

Federico Battiston,¹ Vincenzo Nicosia,¹ and Vito Latora^{1,2}

¹*School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom*

²*Dipartimento di Fisica ed Astronomia, Università di Catania and INFN, I-95123 Catania, Italy*

(Dated: September 15, 2016)

What do societies, the Internet, and the human brain have in common? They are all examples of complex relational systems, whose emerging behaviours are largely determined by the non-trivial networks of interactions among their constituents, namely individuals, computers, or neurons, rather than the properties of the units themselves. In the last two decades, network scientists have proposed models of increasing complexity to better understand real-world systems. Only recently we have realised that multiplexity, i.e. the coexistence of several types of interactions among the constituents of a complex system, is responsible for substantial qualitative and quantitative differences in the type and variety of behaviours that a complex system can exhibit. As a consequence, multilayer and multiplex networks have become a hot topic in complexity science. Here we provide an overview of some of the measures proposed so far to characterise the structure of multiplex networks, and a selection of models aiming at reproducing those structural properties and quantifying their statistical significance. Focusing on a subset of relevant topics, this brief review is a quite comprehensive introduction to the most basic tools for the analysis of multiplex networks observed in the real-world. The wide applicability of multiplex networks as a framework to model complex systems in different fields, from biology to social sciences, and the colloquial tone of the paper will make it an interesting read for researchers working on both theoretical and experimental analysis of networked systems.

I. INTRODUCTION

One of the most intriguing characteristic of complex systems is that most of the collective behaviours they exhibit cannot be predicted from the knowledge of the properties of their elementary constituents. Indeed, in the last two decades network science has shown that in many cases, from biology to economics, the structure of the interactions among the constituents of the system plays a fundamental role in shaping the emergence of complex behaviours, much more important than the role played by the specific properties of the single units of the system [1–4]. And surprisingly, systems as diverse as social networks, transportation systems, cities, and the human brain were shown to share a significant number of features and a comparable structure of interactions [5–10].

Recently, the availability of new data sets, the rediscovery of old ones and the access to more powerful computers, has highlighted the necessity to develop a new framework to represent networks whose units interact through more than just one type of relations. These systems are usually called multiplex networks, and are characterised by the fact that all the connections of a given type are embedded into a distinct layer. A full description of early research on such topic can be found in [11–13]. This Article provides an informal, still comprehensive handbook for the experimental investigation of systems which can be described as multiplex networks. In the first part we provide an overview of the most basic measures to characterise the structure of multiplex networks, focusing on

the properties of nodes, edges, and layers. In the second part we review a few models which can be used to reproduce the empirical patterns observed in real-world multi-layer system, or to assess the statistical significance of such patterns.

II. MEASURES FOR MULTIPLEX NETWORKS

We consider a multiplex network \mathcal{M} consisting of N nodes and M different types of relations, represented by M graphs, or layers. We can fully describe the structure of the system by considering the set of adjacency matrices

$$\mathcal{M} \equiv \mathbf{A} = \{A^{[1]}, \dots, A^{[M]}\}, \quad (1)$$

where $A^{[\alpha]} = \{a_{ij}^{[\alpha]}\}$, with $a_{ij}^{[\alpha]} = 1$ if i and j share a bond of type α and $a_{ij}^{[\alpha]} = 0$ otherwise [14]. When the connections among nodes are weighted, the system can be described by a set of weighted adjacency matrices $\mathbf{W} = \{W^{[1]}, \dots, W^{[M]}\}$, with $W^{[\alpha]} = \{w_{ij}^{[\alpha]}\}$ and $w_{ij}^{[\alpha]}$ the weight of the link between node i and j [15]. This formulation implicitly assumes that each node i consists of M replicas, one at each layer, and that a link can only connect two replicas lying on the same layer. Consequently, although both $\mathbf{A} = \{a_{ij}^{[\alpha]}\}$ and $\mathbf{W} = \{w_{ij}^{[\alpha]}\}$ can be considered as generic order-3 tensors, it is important to stress that node i on layer α and node i on layer β effectively represent the same unit of the system and not two different ones. In other words, the replicas of the same node are identified across layers. Social systems can be naturally cast within this framework, where

different layers can represent for instance different interaction channels among the same nodes (e.g., face-to-face communication, email exchange, online chat, etc.), but the different replicas are just a mathematical representation of the same individual in each of the M contexts.

However, there are cases in which there exists some sort of communication or flow between the replicas of the same node at different layers. A typical example is that of multimodal transportation systems, where nodes are locations and layers represent different transport modalities, e.g. bus, underground, trains, etc. In this case, an accurate modelling of the system has to take into account inter-layer transitions between the replicas of the same node at different layers, and it is more convenient to model the structure of the system through an order-4 tensor $\mathcal{M}_{i,\alpha}^{j,\beta}$ [16]. This formulation makes explicit (and adjustable) the relative importance of intra-layer and inter-layer connections [17, 18]. Unless specified otherwise, in the following we will mostly use the more simple formulation based on order-3 tensors and given in Eq. 1.

A. Node properties

Differently from the traditional single-layer approach, where node properties are described by scalar variables, node features in multiplex networks are naturally described in vectorial terms. As an example, for each node i we consider its total number of connections, or degree, at layer α , i.e. $k_i^{[\alpha]} = \sum_{j \neq i}^N a_{ij}^{[\alpha]}$, and the multilayer degree $\mathbf{k}_i = \{k_i^{[1]}, \dots, k_i^{[M]}\}$. A crucial empirical evidence is that in many multiplex networks not all nodes have connections at all layers. As a consequence, a node i is defined as active on a layer α if it is connected to at least another node at that layer, i.e. if $k_i^{[\alpha]} > 0$. The activity-pattern of each node can be compactly stored into the node-activity vector

$$\mathbf{b}_i = \{b_i^{[1]}, \dots, b_i^{[M]}\}, \quad (2)$$

where $b_i^{[\alpha]} = 1 - \delta_{0, k_i^{[\alpha]}}$, i.e. $b_i^{[\alpha]} = 1$ if node i is active on layer α , and $b_i^{[\alpha]} = 0$ otherwise. The total activity $B_i = \sum_{\alpha=1}^M b_i^{[\alpha]}$ represents the number of layers in which node i is active, with $0 \leq B_i \leq M$ [19]. It has been found that most real-world multiplex networks are characterised by heterogeneous distributions of node activity [19], and it has been shown that such heterogeneity might be responsible for the increased fragility of multiplex networks to random failures [20].

Given a generic vectorial property $\boldsymbol{\xi}_i = \{\xi_i^{[1]}, \dots, \xi_i^{[M]}\}$, it is important to be able to compress the information into meaningful scalar descriptors, especially for systems composed of a large number of layers. A typical way to approach this problem is to consider the first and the second moment of the vector $\boldsymbol{\xi}$, accounting for its mean value $\mu(\boldsymbol{\xi})$ (or, analogously, the sum of its components)

and its variance $\sigma^2(\boldsymbol{\xi})$, or related quantities. In the particular case of the degree, the total number of connections of node i is usually called total or overlapping degree [14]

$$o_i = \sum_{\alpha=1}^N k_i^{[\alpha]} \quad (3)$$

while the heterogeneity of the number of neighbours of node i across the layers can be measured through the multiplex participation coefficient [14]

$$P_i = \frac{M}{M-1} \left[1 - \sum_{\alpha=1}^M \left(\frac{k_i^{[\alpha]}}{o_i} \right)^2 \right], \quad (4)$$

where $P_i = 1$ when the links incident on node i are equally distributed across the layers, and $P_i = 0$ when a node is only active on one layer. We note that similar information about the heterogeneity of the distribution of a node's connections across layers is provided by the Shannon entropy of the degree vector [14]

$$H_i = - \sum_{\alpha=1}^M \frac{k_i^{[\alpha]}}{o_i} \ln \left(\frac{k_i^{[\alpha]}}{o_i} \right). \quad (5)$$

The pair of variables (P_i, o_i) can be used to classify nodes via the so-called multiplex cartography [14], efficiently distinguishing multiplex hubs (high o_i and high P_i), focused hubs (high o_i and low P_i), multiplex leaves (low o_i and high P_i) and focused leaves (low o_i and low P_i).

A remarkable property of real-world networks is the tendency of nodes to form triangles, a phenomenon usually known as transitivity. In single-layer networks, the abundance of triangles is typically measured through the average clustering coefficient C , where $C = \frac{1}{N} \sum_{i=1}^N C_i$ and C_i accounts for the fraction of triads centred on node i which are closed into triangles. In multiplex networks, triads and triangles can effectively extend over more than one layer. We define an m -triad (m -triangle) a triad (triangle) which uses edges from m different layers. It is possible to define two multiplex clustering coefficients to quantify the added value provided to transitivity by the layered structure [14]. For each node i , the first coefficient $C_{i,1}$ is defined as the ratio between the number of 2-triangles with a vertex in i and the number of 1-triads centred in i . In formulas:

$$C_{i,1} = \frac{\sum_{\alpha} \sum_{\alpha' \neq \alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{jm}^{[\alpha']} a_{mi}^{[\alpha]})}{(M-1) \sum_{\alpha} k_i^{[\alpha]} (k_i^{[\alpha]} - 1)} \quad (6)$$

The second multiplex clustering coefficient $C_{i,2}$ is instead defined as the ratio between the number of 3-triangles with node i as a vertex, and the number of 2-triads centred in i . In formulas:

$$C_{i,2} = \frac{\sum_{\alpha} \sum_{\alpha' \neq \alpha} \sum_{\alpha'' \neq \alpha, \alpha'} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{jm}^{[\alpha']} a_{mi}^{[\alpha'']})}{(M-2) \sum_{\alpha} \sum_{\alpha' \neq \alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{mi}^{[\alpha']})}. \quad (7)$$

These two measures are defined respectively for $M \geq 2$ and $M \geq 3$, and are a natural generalisation of clustering coefficient to the case of multiplex networks. A related generalisation of clustering coefficient, based on the order-4 tensorial formulation for multiplex networks, has been suggested in Ref. [21].

Another characteristic properties of real-world networks is the presence of heterogeneity in the relative importance of nodes, as measured by different notions of node centrality [3]. A number of different approaches have been suggested to define and compute the centrality of a node in a multiplex network. A first possibility consists in defining the multiplex centrality as a combination of the centrality scores of each node at the different layers. For instance, starting from the centrality vector of node i , $\mathbf{c}_i = \{c_i^{[1]}, \dots, c_i^{[M]}\}$, one can try to condense the information into a single scalar variable, as is normally done for the degree. However, computing averages of centrality scores across layers is not always meaningful. The first reason is that in general a node can play different roles on different layers, and averaging over layers will only level down such heterogeneities.

The presence of more than one layer allows to define new genuinely multiplex centrality measures in which the role of a node explicitly depends on the structure of the multiplex at all layers. For instance, the authors of Ref. [22] suggested to compute the eigenvector centrality of nodes on each layer α as the normalised eigenvector relative to the largest eigenvalue of

$$\tilde{A}^{[\alpha]} = \sum_{\beta=1}^M i^{[\alpha, \beta]} A^{[\beta]} \quad (8)$$

where $I = \{i^{[\alpha, \beta]}\}$ is a given influence matrix which determines how the centrality of layer α depends on the structure of layer β . In Ref. [14], instead, the authors studied the contributions of the different layers to the centrality of the nodes by varying the coefficients $i^{[\alpha]} \alpha = 1, \dots, M$ of the matrix

$$A' = \sum_{\alpha=1}^M i^{[\alpha]} A^{[\alpha]} \quad (9)$$

which is a convex combination of the adjacency matrices of the layers.

An entire class of node centrality measures can be defined by using the properties of random walks on multiplex networks [23, 24]. A particularly interesting example is that of multiplex PageRank centrality proposed in Ref. [25]. The authors of Ref. [25] considered the case of a two-layer multiplex network and defined the multiplex PageRank of the nodes in layer $\alpha = 2$ as a function of the PageRank scores of the nodes in layer $\alpha = 1$. The main idea of this genuinely multiplex measure is that, especially in social systems, nodes can leverage their centrality in one context, such as personal relationships (represented by layer 1) to gain centrality in another context, e.g. professional relationships (represented by layer 2).

Real-world networks often exhibit the small-world property, meaning that the typical distance between any pair of nodes in the system scales logarithmically with the total number of nodes. An important observation is that not all the nodes of a system are equally important in mediating paths between other nodes, which is the main idea between the concept of node betweenness [3]. In a multiplex system, the reachability of a node might significantly depend on the interplay between different layers. The added value introduced by multiplexity can be measured through the interdependence [26, 27]

$$\lambda_i = \frac{1}{N-1} \sum_{j \neq i} \frac{\psi_{ij}}{\sigma_{ij}}, \quad (10)$$

where ψ_{ij} accounts for the number of shortest paths between i and j that use edges in more than one layer, and σ_{ij} is the total number of shortest paths between i and j . The quantity λ_i takes values in the interval $[0, 1]$, with larger values corresponding to a higher advantage for the reachability of node i provided by the interplay of the different layers. By averaging over all nodes we obtain the interdependence of the multi-layer system $\lambda = 1/N \sum_i \lambda_i$. It is also possible to define a layer interdependence $\lambda^{[\alpha]}$, accounting for the number of shortest paths with at least one link on layer α [28].

Finally, if one represents a multiplex network using an order-4 tensor, an entire class of centrality measures can be obtained as natural extensions to adjacency tensors of the corresponding measures defined on adjacency matrices [30]. For instance, the eigenvector centrality of a node in this formalism can be computed by considering either the eigenvectors of the order-4 tensorial representation of the multiplex or the eigenvectors of the associated supra-adjacency matrix. An interesting application of this class of measures, described in Ref. [29, 30], allows to define the versatility of nodes, assigning higher centrality scores to those nodes which act as bridges among different layers.

B. Layer properties

Similarly to the case of node activity, it is possible to define the activity-vector of each layer α [19] as

$$\mathbf{d}^{[\alpha]} = \{b_1^{[\alpha]}, \dots, b_N^{[\alpha]}\}, \quad (11)$$

where $b_i^{[\alpha]} = 1$ if $k_i^{[\alpha]} > 0$, and $b_i^{[\alpha]} = 0$ otherwise. For each layer α , the total layer activity $N^{[\alpha]} = \sum_{i=1}^N b_i^{[\alpha]}$ describes the total number of nodes with at least one connection in layer α , with $0 \leq N^{[\alpha]} \leq N$. The similarity between the activity-vectors of two layers α and β can be measured by mean of the pairwise multiplexity [19], which accounts for the fraction of nodes of the multiplex which are active on both layers:

$$Q^{[\alpha, \beta]} = \frac{1}{N} \sum_{i=1}^N b_i^{[\alpha]} \cdot b_i^{[\beta]}. \quad (12)$$

In general $0 \leq Q^{[\alpha, \beta]} \leq 1$, with $Q^{[\alpha, \beta]} = 1$ when all nodes are active in both layers, and $Q^{[\alpha, \beta]} = 0$ when no node is active on both layers. The similarity among the patterns of activity in two layers can also be measured through the Hamming distance [19]

$$H^{[\alpha, \beta]} = \frac{\sum_i b_i^{[\alpha]}(1 - b_i^{[\beta]}) + b_i^{[\beta]}(1 - b_i^{[\alpha]})}{\min(N^{[\alpha]} + N^{[\beta]}, N)}. \quad (13)$$

where $H^{[\alpha, \beta]} = 0$ if $\mathbf{d}^{[\alpha]} = \mathbf{d}^{[\beta]}$ and $H^{[\alpha, \beta]} = 1$ if all active nodes are active in no more than one layer. It has been suggested that real multiplex networks are normally characterised by heterogeneous distributions of layer activity and of pairwise multiplexity [19].

Another interesting property observed in real multiplex networks is the presence of correlations between the degrees of the same node at different layers. This is normally signalled by the fact that the probability $P(k^{[\alpha]} = k_1, k^{[\beta]} = k_2)$ to find a node with degree k_1 on layer α and degree k_2 on layer β does not factorise in the product $P^{[\alpha]}(k)P^{[\beta]}(k)$ of the degree distributions of the two layers. In general, given two layers α and β and a generic node property ξ_i , the correlation between $\xi_i^{[\alpha]}$ and $\xi_i^{[\beta]}$ can be computed using the rank correlation coefficient [19]:

$$\rho^{[\alpha, \beta]} = \frac{\sum_i (R_i^{[\alpha]} - \overline{R^{[\alpha]}})(R_i^{[\beta]} - \overline{R^{[\beta]}})}{\sqrt{\sum_i (R_i^{[\alpha]} - \overline{R^{[\alpha]}})^2 \sum_j (R_j^{[\beta]} - \overline{R^{[\beta]}})^2}} \quad (14)$$

where $R_i^{[\alpha]}$ is the rank of node i at layer α induced by the property ξ . When the property of interest is the degree, it makes sense to define the quantity:

$$\overline{k^{[\beta]}}(k^{[\alpha]}) = \sum_{k^{[\beta]}} k^{[\beta]} P(k^{[\beta]} | k^{[\alpha]}) \quad (15)$$

that is the average degree at layer β of a node having degree $k^{[\alpha]}$ at layer α , and is the multiplex homologous of the nearest-neighbours average degree function $k_{nn}(k)$ traditionally used to quantify degree-degree correlations in single-layer graphs [31]. An increasing (decreasing) trend in $\overline{k^{[\beta]}}(k^{[\alpha]})$ will signal the presence of positive (negative) inter-layer degree correlations between layer α and layer β .

The authors of Ref. [32] proposed to quantify inter-layer degree correlations by using the pairwise mutual information between the degree sequences of the two layers:

$$I^{[\alpha, \beta]} = \sum_{k^{[\alpha]}} \sum_{k^{[\beta]}} P(k^{[\alpha]}, k^{[\beta]}) \log \frac{P(k^{[\alpha]}, k^{[\beta]})}{P(k^{[\alpha]})P(k^{[\beta]})} \quad (16)$$

which is maximal when the degree sequences $\{k_i^{[\alpha]}\}$ and $\{k_i^{[\beta]}\}$ are perfectly correlated (or perfectly anti-correlated), and minimal when they are uncorrelated.

An equivalent set of quantities to measure the inter-layer assortativity has been defined for the order-4 tensorial formulation in Ref. [33].

An analysis of the spectral properties of multiplex networks can be found in [34, 35].

A fundamental research question in the field of multiplex networks is to assess whether the presence of more than one interaction layer indeed provides more information about the structure of a system compared to a classical single-layer network representation. In particular, it is interesting to quantify how much information is lost (if any at all) when we aggregate some or all the layers of a multiplex network to obtain a lower-dimensional representation. The authors of Ref. [36] tackled the problem of multiplex reducibility by drawing on an existing formal parallel between density operators of quantum systems and Laplacian matrices of graphs, and extending the concept of Von Neumann entropy of a graph to the case of multiplex networks. They proposed a greedy procedure, based on the estimation of the quantum Jensen-Shannon divergence between layers, which allows to successively aggregate the most redundant layers of a multiplex and to obtain a more compact representation which uses the minimal number of layers while maximising the distinguishability between the multiplex and the single-layer representation of the same system. An interesting result of the paper is that different multilayer systems allow different levels of reducibility, with man-made systems being the least reducible and biological and social systems showing the highest levels of redundancy [36].

Interestingly, many real-world multiplex networks are far from being random combinations of the different layers, but their structures were found to be determined by hidden geometric correlations [37].

C. Edge properties

Due to the presence of multiple layers, a pair of nodes (i, j) can be connected through several edges. Given two layers α and β , the edge overlap of the pair (i, j) [14, 38] is defined as

$$o_{ij}^{[\alpha, \beta]} = \frac{a_{ij}^{[\alpha]} + a_{ij}^{[\beta]}}{2}, \quad (17)$$

where $o_{ij}^{[\alpha, \beta]} = 1$ if i and j are connected at both layers, $o_{ij}^{[\alpha, \beta]} = 1/2$ if they are connected at one layer only, and $o_{ij}^{[\alpha, \beta]} = 0$ if the two nodes are not connected. For a generic number of layers M , the edge overlap is defined as

$$o_{ij} = \frac{1}{M} \sum_{\alpha} a_{ij}^{[\alpha]}. \quad (18)$$

This measure can be easily extended to the whole network as

$$o = \frac{2}{N(N-1)} \sum_{i, j \neq i} o_{ij}, \quad (19)$$

where the average is computed over all possible pairs of nodes [14], or instead as

$$\omega = \frac{\sum_i \sum_{j>i} a_{ij}^{[\alpha]}}{M \sum_i \sum_j 1 - \delta_{0, \sum_\alpha a_{ij}^{[\alpha]}}} \quad (20)$$

where the average is restricted to the pairs of nodes which share at least one edge [32]. Alternative definitions for the local edge overlap on a node i and the total overlap of two layers are suggested in [38] and respectively read

$$\tilde{o}_i^{[\alpha, \beta]} = \sum_{j=1}^N \tilde{o}_{ij}^{[\alpha, \beta]} = \sum_{j=1}^N a_{ij}^{[\alpha]} a_{ij}^{[\beta]}, \quad (21)$$

and

$$\tilde{o}^{[\alpha, \beta]} = \sum_{i<j} a_{ij}^{[\alpha]} a_{ij}^{[\beta]}, \quad (22)$$

where $\tilde{o}_{ij}^{[\alpha, \beta]} = 1$ when both layers have a link between i and j and $\tilde{o}_{ij}^{[\alpha, \beta]} = 0$ otherwise. In the same spirit, a similar measure of edge correlations is the so-called multiplexity [39], defined as

$$m^{[\alpha, \beta]} = \frac{2 \sum_{i<j} \min(a_{ij}^{[\alpha]}, a_{ij}^{[\beta]})}{K^{[\alpha]} + K^{[\beta]}} \quad (23)$$

where $K^{[\alpha]}$ ($K^{[\beta]}$) is the total number of edges at layer α (β). Notice that $m^{[\alpha, \beta]}$ takes values in the range $[0, 1]$.

A somehow dual quantity is the so-called edge intersection index,

$$INT = M \frac{\sum_{i,j=1}^N \min_\alpha \{a_{ij}^{[1]}, a_{ij}^{[2]}, \dots, a_{ij}^{[M]}\}}{\sum_{\alpha=1}^M \sum_{i,j=1}^N a_{ij}^{[\alpha]}} \quad (24)$$

which measures the probability of finding a pair of nodes that is connected by an edge on all the M layers of the multiplex [36].

An alternative characterisation of edge correlations can be based on the conditional probability to find a link at layer α given the presence of an edge between the same nodes at layer β [14]

$$P(a_{ij}^{[\alpha]} | a_{ij}^{[\beta]}) = \frac{\sum_{ij} a_{ij}^{[\alpha]} a_{ij}^{[\beta]}}{\sum_{ij} a_{ij}^{[\beta]}} \quad (25)$$

If layer β has weighted edges, it is also possible to look at the conditional probability $P^w(a_{ij}^{[\alpha]} | w_{ij}^{[\beta]})$ to have a link at layer α given its weight on layer β . If P^w shows an increasing trend as a function of w , this phenomenon goes under the name of edge reinforcement, since a stronger link on one layer implies a higher chance to find the same edge on a different layer [14].

D. Mesoscale properties

Complex networks are usually characterised by non-trivial structural patterns not only at the level of single-node properties but also, and more importantly, at the level of sub-graphs. A lot of attention has been devoted to the analysis of statistically significant sub-graphs in single-layer networks, also known as motifs. It has been found that a few specific sub-graphs are over-represented in real systems compared to their abundance in equivalent networks obtained by randomising the original graph [40, 41]. Due to the additional level of richness provided by the layered structure of multiplex networks, the multilink, i.e. the organisation of the edges between the same pair of nodes (i, j) across the M layers, is the most basic motif [15, 38]. Similarly, m-triads and m-triangles used for the definition of node clustering coefficients are multiplex motifs [14]. The problem of isomorphisms in multi-layer is studied in Ref. [42]. A general classification in three levels of higher-order motifs is presented in Ref. [43]: at the first level connected subgraphs are distinguished according to their number of nodes; at the second level different patterns are classified on the aggregated network and eventually, at the third level, the exact multiplex connectivity pattern is identified. Such approach has been used for the analysis of two-layer networks based on structural and functional connectivity in the human brain [43].

Another remarkable feature of networked systems is the tendency of their units to cluster together in tightly-knit groups, giving rise to non-trivial community structures. Communities are also observed in multiplex networks, even if there is not to date an agreed definition of what a multiplex community is [44]. Some of the efforts in the characterisation of the communities of a multiplex have been focused on the quantification of the similarities in the community structure observed at different layers. In general, given two layers α and β and their partitions in communities \mathcal{P}_α and \mathcal{P}_β , their similarity can be measured through the normalised mutual information (NMI) [45, 46]

$$NMI(\mathcal{P}_\alpha, \mathcal{P}_\beta) = \frac{-2 \sum_{m=1}^{M_\alpha} \sum_{m'=1}^{M_\beta} N_{mm'} \log \left(\frac{N_{mm'} N}{N_m N_{m'}} \right)}{\sum_{m=1}^{M_\alpha} N_m \log \left(\frac{N_m}{N} \right) + \sum_{m'=1}^{M_\beta} N_{m'} \log \left(\frac{N_{m'}}{N} \right)} \quad (26)$$

where $N_{mm'}$ is the number of nodes in common between community m of partition \mathcal{P}_α and community m' of partition \mathcal{P}_β , while N_m and $N_{m'}$ are respectively the number nodes in the two communities m and m' . We note that such measure was originally suggested to compare the community structure obtained on the same single-layer networks from different algorithms. A different similarity measure is suggested in Ref. [47], in terms of the possibility to infer the community structure at layer α using information about the community structure at layer β .

The information about the decomposition in communities of different layers can be combined together to de-

fine a multilayer partition in communities. A classical approach is that described in Ref. [44], which proposed a generalisation of the concept of modularity to multiplex, multi-slice, and temporal networks. Among the genuinely multiplex methods to extract the community decomposition of a system we find particularly interesting the approach proposed in Ref. [45], which extends the Infomap algorithm [48], based on the minimisation of the description length of a partition in communities, to the case of multiplex networks.

III. MODELS OF MULTIPLEX NETWORKS

The characterisation of the structure of a network is normally accompanied by a modelling effort aiming at quantifying how special or peculiar are the observed patterns, e.g. in terms of how probable is to find them in an appropriately chosen family of random graphs, and which are the mechanisms that determine their appearance. In the following we review two classes of models of multiplex networks, namely static random graph models and growing networks.

A. Microcanonical and Canonical Ensembles

A standard approach to study the structure of a given network is to quantify how probable is to observe a network with similar properties in an appropriately defined ensemble of random graphs whose elements satisfy certain constraints. For instance, it is a well-known fact that graphs with power-law degree distributions are extremely rare in the classical Erdős-Renyi random graph ensemble, where each pair of nodes are connected with a given probability p . As a consequence, the hypothesis that power-law degree distributions arise as a result of a uniform distribution of edges across the nodes can be safely rejected, and we can conclude that some other mechanism should be at work in the formation of graphs with heterogeneous degree sequences.

An ensemble of graphs is determined by a set of constraints that its elements should satisfy. According to the type of constraints, we can identify at least two classes of random network ensembles, namely canonical ensembles, where each graph of the ensemble satisfies the set of constraints on average (soft constraints), and microcanonical ensembles, where each graph satisfies all the constraints exactly (hard constraints). It is possible to define a sequence of canonical and microcanonical ensembles of multiplex networks [38], where the constraints are just the average degree at each of the M layers, or the degree distribution of each layer, or the degree distribution together with the distribution of edge overlap, and so on.

Each multiplex networks ensemble is defined by providing the probability $P(\mathbf{A})$ for each of the possible configuration of multiplex networks \mathbf{A} which satisfy the con-

straints. Starting from $P(\mathbf{A})$, the Shannon entropy of the ensemble is defined as

$$S = - \sum_{\mathcal{M}} P(\mathcal{M}) \ln P(\mathcal{M}) \quad (27)$$

[38].

For the special case of uncorrelated multiplex networks, we have

$$\langle a_{ij}^{[\alpha]} a_{ij}^{[\beta]} \rangle = \langle a_{ij}^{[\alpha]} \rangle \langle a_{ij}^{[\beta]} \rangle \quad (28)$$

the probability $P(\mathcal{M}) \equiv P(\mathbf{A})$ can be factorised into the probability of observing each single layer, i.e.

$$P(\mathbf{A}) = \prod_{\alpha=1}^M P^{[\alpha]}(A^{[\alpha]}). \quad (29)$$

In this particular case, the entropy of the multiplex ensemble reads

$$S = \sum_{\alpha=1}^M S^{[\alpha]} = - \sum_{\alpha=1}^M P^{[\alpha]}(A^{[\alpha]}) \ln(P^{[\alpha]}(A^{[\alpha]})). \quad (30)$$

In the following we focus on the canonical - indicated by C - and microcanonical - denoted by M - ensembles of multiplex networks. Let us assume that we have T soft constraints such that

$$\sum_{\mathbf{A}} P(\mathcal{M}) F_{\mu}(\mathcal{M}) = C_{\mu} \quad (31)$$

where $\mu = 1, \dots, T$, and $F_{\mu}(\mathcal{M})$ describes how such constraints are imposed on the network, such as the degree of each node of the network at each layer α , or the total number of edges $K^{[\alpha]}$ for $\alpha = 1, \dots, M$. The probability $P_C(\mathcal{M})$ of observing the multiplex \mathcal{M} can be obtained by maximising the entropy S under the given set of constraints. By solving the optimisation problem one obtains:

$$P_C(\mathcal{M}) = \frac{1}{Z_C} \exp \left[- \sum_{\mu} \lambda_{\mu} F_{\mu}(\mathcal{M}) \right] \quad (32)$$

where Z_C is the partition function of the canonical multiplex ensemble and the values of the Lagrangian multipliers λ_{μ} are obtained by satisfying Eq.31 imposing such functional form for $P_C(\mathcal{M})$. In the canonical multiplex ensembles we have:

$$S = \sum_{\mu} \lambda_{\mu} C_{\mu} + \ln Z_C. \quad (33)$$

Conversely, in the microcanonical multiplex ensemble each multiplex configuration compatible with the hard constraints has the same probability

$$P_M(\mathcal{M}) = \frac{1}{Z_M} \prod_{\mu=1}^T \delta[F_{\mu}(\mathcal{M}), C_{\mu}] \quad (34)$$

where δ is the Kronecker delta function and $Z_M = \sum_{\mathbf{A}} \prod_{\mu=1}^T \delta[F_\mu(\mathcal{M}), C_\mu]$ is the microcanonical partition function, accounting for the number of multiplex networks satisfying the T hard constraints $F_\mu(\mathcal{M}) = C_\mu$. By defining the entropy of these ensembles as $N\Sigma$, such entropy reads

$$N\Sigma = \sum_{\mathbf{A}} P_M(\mathcal{M}) \ln P_M(\mathcal{M}) = \ln Z_M \quad (35)$$

where Σ is the Gibbs entropy of the multiplex ensemble. It can be shown that the Gibbs entropy Σ is related to the corresponding Shannon entropy S by $N\Sigma = S - N\Omega$, where Ω is the logarithm of the probability that in the related canonical multiplex ensemble the hard constraints $F_\mu(\mathcal{M})$ are satisfied.

The author of Ref. [38] provided an exhaustive explanation of how the entropy and the partition function can be computed in different classes of multiplex networks with increasingly stringent sets of constraints, both for the canonical and for the microcanonical ensembles. The same approach has been generalised to a number of more complicated structures, including spatial multiplex networks [49] and multiplex networks with heterogeneous activities of the nodes [20]. The Authors of Ref. [50] study the canonical ensemble of the overlapping networks generated by merging different layers, where information on the connection between nodes is only accessible at the aggregated level.

B. Models of node and layer activity

The concept of node and layer activity is peculiar to multilayer networks, and it is interesting to assess whether simple models can give account for the observed heterogeneous distributions of node and layer activities. In the following we provide a brief review of some null models proposed so far to quantify the peculiarity of given distributions of node and layer activities.

Let us consider two layers α and β with $N^{[\alpha]}$ and $N^{[\beta]}$ active nodes respectively. If initially the two layers have no active nodes and we then sample uniformly at random from $\{1, 2, 3, \dots, N\}$ $N^{[\alpha]}$ nodes on layer α and $N^{[\beta]}$ nodes on layer β and we activate them, then the probability that m of them are active at both layers follows a hypergeometric distribution

$$p(m; N, N^{[\alpha]}, N^{[\beta]}) = \frac{\binom{N^{[\alpha]}}{m} \binom{N - N^{[\alpha]}}{N^{[\beta]} - m}}{\binom{N}{N^{[\beta]}}}, \quad (36)$$

according to which the expected number of nodes active at both layers is equal to $N^{[\alpha]}N^{[\beta]}/N$, the expected pairwise multiplexity is

$$\tilde{Q}^{[\alpha, \beta]} = \frac{N^{[\alpha]}N^{[\beta]}}{N^2} \quad (37)$$

and the expected Hamming distance reads

$$\tilde{H}^{[\alpha, \beta]} = \frac{\sum_{m=0}^{N^{[\beta]}} (N^{[\alpha]} + N^{[\beta]} - 2m) \times p(m; N, N^{[\alpha]}, N^{[\beta]})}{\min(N, N^{[\alpha]} + N^{[\beta]})} \quad (38)$$

This is the simplest model of node activation and is known as the hypergeometric model [19]. However, the authors of Ref. [19] have shown that the distribution of pairwise multiplexity and pairwise Hamming distance in real-world multiplex networks is not compatible with those given in Eq. 37 and Eq. 38.

Let us now consider the problem of constructing a multiplex networks with a fixed number of layers M , a fixed number of nodes N which are active on at least one of the M layers, and where each node i has an assigned node activity B_i , which is for instance set equal to the node activity observed in the real network. By sampling for each node one of the $\binom{M}{B_i}$ vectors of node-activity with B_i non-zero entries, the distribution of the total node activity of the original system is kept fixed, whereas the correlations in the layer activity and the distribution of the node-activity vectors are not preserved. Moreover, in such a model all layers have the same expected number of active nodes:

$$\tilde{N}^{[\alpha]} = \frac{1}{M} \sum_i B_i. \quad (39)$$

This model is known as the multi-activity deterministic model [19]. A variation of the model is constructed by activating node i in each layer α with probability $\bar{B}_i = B_i/M$, so that the expected activity of each layer stays the same but the original node-activity distribution is not preserved. This model is known as the multi-activity stochastic model [19].

Finally, it is possible to construct a model for a 2-layer multiplex network where the degree distributions of each layer is kept fixed, and where one can control the edge overlap ω by rewiring a certain fraction r of the edges. The model was introduced in Ref. [51]. For simplicity, let us assume that the two layers have the same number of edges $K^{[\alpha]} = K^{[\beta]} = K$. If we start from two identical networks, we have maximum edge overlap $\omega = 1$. If we now keep fixed the structure on one of the two layers, and rewire one of the edges of the other layer, the number of links present in both layers decreases by one unit, while the number of those present in only one of the two layers increases by two units. Consequently, if we rewire a fraction r of the K edges of the second layer in such a way that each rewire decreases the number of edges existing on both layers, we obtain:

$$\omega = \frac{(1-r)K}{(1+r)K} = \frac{(1-r)}{(1+r)} \quad (40)$$

By inverting such relation, we find that a given overlap ω corresponds to a rewire r equal to $r = (1-\omega)/(1+\omega)$. In practice, this model allows to obtain a prescribed value of edge overlap by rewiring a certain fraction r of the edges in one of the two layers.

C. Growth models of multiplex networks

In this section we review a few growth models for multiplex networks. The most simple example of this class is a model of layer-growth, aimed at explaining the fat-tail distribution of layer activity observed in empirical data [19]. The model works as follows. We start at time $t_0 = 0$ with a multiplex with M_0 layers and N nodes. At each time t , a new layer α joins the network with $N^{[\alpha]}$ nodes to be activated, where $N^{[\alpha]}$ can be observed from the data-set we are attempting to reproduce. Each node i has then a probability to be active on that layer at time t equal to:

$$p_i(t) = A + B_i(t), \quad (41)$$

where $B_i(t)$ is the number of layers where node i is already active and $A > 0$ is a constant that allows the activation of nodes not yet active in the multiplex. When the number of layers in the model increases, the distribution of layer activity $P(N^{[\alpha]})$ approaches a power law.

Another important class of growth models is that where not layers, but individual nodes join sequentially the network, for instance by connecting to preexisting vertices on possibly different layers. In such regard, it is clear that the specific shape of the attachment function determines the long-term statistical properties of the final multiplex graph. In single-layer networks, a particularly well-studied case is the so-called preferential attachment, where nodes choose to attach to older vertices depending on a function (in the simplest case linear) of their degree k . In a multiplex network the degree of each node j is a vector and the probability $\Pi_{i \rightarrow j}^{[\alpha]}$ that a new node i attaches to j on a given layer α in general depends on all its components. In formula:

$$\Pi_{i \rightarrow j}^{[\alpha]} = \frac{F_j^{[\alpha]}(\mathbf{k}_j)}{\sum_{\ell} F_{\ell}^{[\alpha]}(\mathbf{k}_{\ell})} \quad (42)$$

The most simple class of preferential attachment models is obtained by considering linear attachment kernels, i.e. by setting $F_j^{[\alpha]}$ as a convex combination of the degrees of node j at all the layers [27, 52]. The interesting result is that linear attachment kernels produce multiplex networks whose layers have power-law degree distributions, but where inter-layer degree correlations are always positive, meaning that a hub on one layer is also a hub on the other layer as well. This is due to the fact that the expected final degree of a node on a certain layer is determined solely by the time at which it joins the network [27]. A generalisation of the closed-form solutions for the joint degree distribution of heterogeneously growing multiplex networks with arbitrary number of layers and arbitrary times can be found in [53].

A more interesting class of multiplex networks is obtained by considering non-linear attachment kernels. The authors of Ref. [54] started from the case of multiplex

networks with two layers, using the attachment kernel:

$$F_j^{[1]} \propto \left(k_j^{[1]}\right)^{\alpha} \left(k_j^{[2]}\right)^{\beta} \quad (43)$$

where $\alpha, \beta \in \mathbb{R}$. By tuning the relative values of the exponents α and β , one can obtain multiplex networks where each layer has either an exponential, a power-law, or a condensed degree distribution (where super-hubs with extensive degrees appear). Moreover, with non-linearity it is possible to get not only positive inter-layer degree correlations as in the case of the linear model, but also null and negative correlations that have been observed in real-world systems [19]. In the same work the authors suggested several possible generalisations of the model to the case of multiplex networks with $M \leq 2$ layers. An interesting model of multiplex network growth which takes into account weighted links, aimed at reproducing the structure of some layered social networks, can be found in Ref. [55].

D. Models of multiplex communities

Simple preferential attachment models, while able to reproduce some empirical patterns such as inter-layer degree correlations, do not allow to construct multiplex networks with strong community structure. More sophisticated models able to produce tunable intra-layer and inter-layer community structure have been suggested, based on intra-layer and inter-layer triadic closure mechanisms on 2-layer multiplexes [46]. In that model a node i arrives and selects one of the layers at random, and a node n_1 in that layer as its first neighbour. The following $m - 1$ links on the same layer will be to a neighbour of n_1 with probability p , or to a uniformly sampled node in the same layer with probability $1 - p$. Once m links have been created on the first layer, node i starts creating links on the other layer. In particular, the first edge on the other layer is created with probability p^* to the same node n_1 , and at to a node sampled uniformly at random with probability $1 - p^*$. The remaining links on the second layer are placed again with probability p and $1 - p$. In this model, the value of the parameter p determines the strength of communities on each layer, with higher values of p corresponding to tighter communities, while p^* tunes the extent of overlap (i.e., number of shared nodes) between communities in different layers. Interestingly, the model was able to reproduce some of salient characteristics of multiplex collaboration networks.

We note here that multiplex networks with given community structure can also be generated through stochastic block models [56], and that in some cases aggregated networks can be better fitted by multilayer block models, hinting at the existence of different levels in the considered data [57].

IV. CONCLUSIONS

Networks are responsible for the emergence of a variety of complex behaviours in social, economical, technological and biological systems, and multilayer networks are the last frontier of research in this field. The theory of multiplex networks has already proven quite successful in modelling the structure of intrinsically multidimensional relational systems, showing at the same time that

the presence of more than a single type of interaction is responsible for new levels of complexity. The advances made in this field in the last few years are definitely encouraging, and there is still a lot of open problems to address in depth and many questions still waiting to be asked. We strongly believe that multiplex networks are an extremely active and interesting area of research, and we really hope that this brief review will contribute to spur the curiosity of researchers who are interested in studying the structure of real-world systems.

-
- [1] R. Albert, A.L. Barabási, *Reviews of Modern Physics* **74**, 47 (2002)
 - [2] M. E. J. Newman, *SIAM Review* **45**, 167-256 (2003)
 - [3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D. U. Hwang, *Physics Reports* **424**, 175-308 (2006)
 - [4] M. E. J. Newman, *Networks: an Introduction* (Oxford University Press, Oxford, UK) (2010)
 - [5] S. Strogatz, *Nature* **410**, 268-276 (2001)
 - [6] A. Arenas et al., *Physics Reports* **469** (3), 93-153 (2008)
 - [7] E. Bullmore, O. Sporns, *Nature Reviews on Neuroscience* **10**, 186-198 (2009)
 - [8] C. Castellano, S. Fortunato, V. Loreto, *Reviews of Modern Physics* **81**, 591-646 (2009)
 - [9] M. Barthélemy, *Physics Reports* **499**, 1-101 (2011)
 - [10] F. De Vico Fallani, J. Richiardi, M. Chavez, S. Achard, *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **369** (2014)
 - [11] S. Boccaletti et al., *Physics Reports* **544** 1 1-122 (2014)
 - [12] M. Kivela et al., *Journal of Complex Networks* **2** (3) 203 - 271 (2014)
 - [13] K.-M. Lee, B. Mina, K.-I. Goh, *Eur. Phys. J. B* **88**, 48 (2015)
 - [14] F. Battiston, V. Nicosia, V. Latora, *Physical Review E* **89** 032804 (2014)
 - [15] G. Menichetti et al., *PLoS ONE* **9** (6): e97857. doi:10.1371/journal.pone.0097857 (2014)
 - [16] M. De Domenico et al., *Physical Review X* **3** (4), 041022 (2013)
 - [17] S. Gómez et al., *Physical Review Letters* **110**, 028701 (2013)
 - [18] F. Radicchi, A. Arenas, *Nature Physics*, **9**, 717 (2013)
 - [19] V. Nicosia, V. Latora, *Physical Review E* **92** 032805 (2015)
 - [20] D. Cellai, G. Bianconi, *Physical Review E* **93**, 032302 (2016)
 - [21] E. Cozzo et al., *New Journal of Physics* **17** (7), 073029 (2015)
 - [22] L. Solá et al., *Chaos: An Interdisciplinary Journal of Nonlinear Science* **23** (3), 033131 (2013)
 - [23] F. Battiston, V. Nicosia, V. Latora, *New Journal of Physics* **18** (4), 043035 (2016)
 - [24] A. Solé-Ribalta, M. De Domenico, S. Gómez and A. Arenas, *Physica D* **323**, 73-79 (2016)
 - [25] A. Halu, R. J. Mondragón, P. Panzarasa, G. Bianconi, *PLoS ONE*, **8** (10): e78293. doi:10.1371/journal.pone.0078293 (2013)
 - [26] M. G. Morris, M. Barthelemy, *Physical Review Letters* **109** 12 128703 (2012)
 - [27] V. Nicosia, G. Bianconi, V. Latora, M. Barthelemy, *Physical Review Letters* **111** 058701 (2013)
 - [28] A. Aleta, S. Meloni, Y. Moreno, arXiv preprint, arXiv:1607.00072 (2016)
 - [29] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, A. Arenas, *Nature Communications* **6** (2015)
 - [30] A. Solé-Ribalta, M. De Domenico, S. Gómez, A. Arenas, *Proceedings of the 2014 ACM conference on Web Science* 149-155 (2014)
 - [31] R. Pastor-Satorras, A. Vazquez, A. Vespignani, *Physical Review Letters* **87**, 258701 (2001)
 - [32] L. Lacasa, V. Nicosia, V. Latora, *Scientific reports* **5** (2015)
 - [33] G. Ferraz de Arruda, E. Cozzo, Y. Moreno, F. A. Rodrigues, arXiv preprint, arXiv:1507.04550 (2015)
 - [34] A. Solé-Ribalta et al., *Physical Review E* **88** (3), 032807 (2013)
 - [35] R.J. Sánchez-García, E. Cozzo, Y. Moreno, *Physical Review E* **89** (5), 052815 (2014)
 - [36] M. De Domenico, V. Nicosia, A. Arenas, V. Latora, *Nature Communications* **6** (2015)
 - [37] K.-K. Kleineberg, M. Boguna, M. Angeles Serrano, F. Papadopoulos, *Nature Physics* doi:10.1038/nphys3812 (2016)
 - [38] G. Bianconi, *Physical Review E* **87** 6 062806 (2013)
 - [39] V. Gemmetto, D. Garlaschelli, *Scientific Reports* **5** 9120 (2015)
 - [40] R. Milo et al., *Science* **298**, 824-827 (2002)
 - [41] R. Milo et al., *Science* **303**, 1538-1542 (2004)
 - [42] M. Kivela, M.A. Porter, arXiv preprint arXiv:1506.00508 (2015)
 - [43] F. Battiston, M. Chavez, V. Nicosia, V. Latora, arXiv preprint arXiv:1606.09115 (2016)
 - [44] P. J. Mucha et al., *Science* **328**, 876-878 (2010)
 - [45] M De Domenico, A Lancichinetti, A Arenas, M Rosvall, *Physical Review X* **5** (1), 011027 (2015)
 - [46] F. Battiston, J. Iacovacci, V. Nicosia, G. Bianconi, V. Latora, *PLOS ONE* **11** (1) e0147451 (2016)
 - [47] J. Iacovacci, Z. Wu, G. Bianconi, *Physical Review E* **92** (4), 042806 (2015)
 - [48] M. Rosvall, C. T. Bergstrom, *Proceedings of the National Academy of Sciences of the U.S.A.* **105**, 1118-1123 (2008)
 - [49] A. Halu, S. Mukherjee, G. Bianconi, *Physical Review E* **89**, 012806 (2014)
 - [50] O. Sagarra, C. J. Perez Vicente, A. Diaz-Guilera, *Physical Review E* **92** (5), 052816 (2015)
 - [51] M. Diakonova, V. Nicosia, V. Latora, M. San Miguel, *New J. Phys.* **18** 023010 (2016)

- [52] J. Y. Kim and K.-I. Goh, Physical Review Letters **111**, 058702 (2013).
- [53] N. Momeni, B. Fotouhi, Physical Review E **92**, 062812 (2015)
- [54] V. Nicosia, G. Bianconi, V. Latora, M. Barthelemy, Physical Review E **90**, 042807 (2014)
- [55] Y. Murase, J. Török, H.-H. Jo, K. Kaski, J. Kertész, Physical Review E **90**, 052810 (2014)
- [56] T. P. Peixoto, Physical Review E **92**, 042807 (2015)
- [57] T. Valles-Catala, F. A. Massucci, R. Guimera, M. Sales-Pardo, Physical Review X **6**, 011036 (2016)